

DEMONSTRATIVE PRONOUNS IN BALTO-SLAVIC,
ARMENIAN, AND TOCHARIAN

FREDERIK KORTLANDT

1. A reconstruction of the Balto-Slavic demonstrative pronouns must be based on a comparison of the Prussian (Pr.), Lithuanian (Li.), and Old Slavic (Sl.) evidence. Such a comparison turns out to be more complicated than it might appear at first sight.

The PIE. demonstrative pronoun *so*, *sā*, *tod* is reflected in Sl. *rz*, Li. *tàs*, and Pr. *stas*. Van Wijk has conclusively argued that the Prussian forms arose from a contamination of *sa* and *ta-* (1918: 111). It follows that the suppletive nom.sg. forms *so* and *sā* had been preserved in the Balto-Slavic proto-language.

Similarly, the stem *ono-* (*ano-*) provides suppletive nominative forms for the pronominal stem *i(o)-* in both Slavic and Old Lithuanian (cf. Van Wijk 1918: 116). Since the nom.pl. forms are probably secondary (*ibidem*: 125), we can assume for Balto-Slavic the same distribution here as in the case of *so*, *sā*, *to-*.

Consider now the following normalized paradigm of the Pr. 3rd person pronoun as it appears in the *Enchiridion*:

- sg. nom. m. *tāns*, f. *tennā*.
 gen. m. *tennessei*, *tennēison*.
 dat. m. *tennesmu*, f. *tennēi*.
 acc. mf. *tennan*, *din*.
 pl. nom. m. *tennei*, *dei*.
 gen. mf. *tennēison*.
 dat. m. *tennēimans*.
 acc. mf. *tennans*, *dins*.

The acc. forms *tennan* and *tennans* are evidently based on the stem *ono-* (*ano-*), which underlies the nom.sg. forms *tāns* and *tennā*, whereas the other case forms betray an alternating stem *e-*, *ei-*, *i-*. This state of affairs is more archaic than the one attested in the flexion of the Li. and Sl. 3rd person pronoun, where we find an alternation between *i-* and *io-*.

The problem becomes even more complicated if we take the flexion of the Pr. demonstrative pronoun *stas* into account. If we disregard its use as an article, the paradigm can be normalized as follows:

- sg. nom. m. *stas*, f. *stai*, n. *sta*.
 gen. m. *stessei*, f. *stessies*.
 dat. m. *stesmu*, f. *stessiei*.
 acc. mf. *stan*.
 pl. nom. mf. *stai*.
 gen. mf. *stēison*.
 dat. mf. *stēimans*.
 acc. mf. *stans*.

For the article the paradigm must be modified as follows:

- sg. gen. m. *stessei*, *stēisei*, *stēison*,
 f. *stesses*, *stessei*, *stēisei*, *stēison*.
 dat. m. *stesmu*, *stēismu*,
 f. *stessei*, *stēisei*, *stesmu*, *stēismu*.

If we assume that only the first form of every line is original, we arrive at a paradigm which differs from the pronoun in the feminine gen. and dat.sg. forms only. Elsewhere I have argued that this difference is probably old and that we have to compare the article with Skt. *asyās*, *asyái* < *esiās*, *-āi* and the pronoun with *táśyās*, *táśyai* < *tosīās*, *-āi* (1982: 9). This leads us to the reconstruction of three types of demonstrative pronoun in order to account for the Prussian material:

- an anaphoric pronoun *to-* + *ono-* (*ano-*) + *e/i-*,
- a weak demonstrative *so* + *to-* + *e/i-*,
- a strong demonstrative *so* + *to-*.

2. Apart from the pronouns mentioned so far, there is an unmistakable correspondence between Sl. *čъ*, Li. *šīs*, and Pr. *schis*, all of which designate *hic* deixis. Here again, the details are such that they preclude a straightforward reconstruction of the Balto-Slavic paradigm. The Slavic word is inflected as a soft stem in spite of the fact that the stem was undoubtedly hard, as is clear from the West Slavic reflex *s-*, not *š-*. It follows that the formation of the Slavic paradigm was probably posterior to the umlaut and that we have to assume that the alternating stem *e-*, *ei-*, *i-* was preserved in early Slavic. This hypothesis also accounts for the aberrant masculine nom.pl. form *чнн* (cf. Vaillant 1958: 383). The gen. and dat.sg. forms *чєро*, *чєрѣ* and *чємовъ*, *чєм* suggest that the formation of the new paradigm can be dated between the umlaut and the rise of prothetic *j-* before *e-*, i.e. to what I have called the Early Middle Slavic period (1979: 263). If it had been earlier or later, we would expect *še-* or *sije-* rather than *se-*.

The flexion of Pr. *schis* is particularly difficult to establish. I am inclined to start from the following normalized paradigm:

sg. nom. m. *schis*.

gen. m. *schissei*.

dat. m. *schismu*.

acc. mf. *schan, schien*.

pl. nom. m. *schai*.

gen. mf. *schieison*.

acc. mf. *schans, schiens*.

It is attractive to regard *schan*, *schans* and *schien*, *schiens* as original masc. and fem. forms, respectively, though the material is too small to substantiate this hypothesis. If it is correct, the paradigm contains three different stems: *ši-* in the nom.sg. form and in the case forms which have *-e-* in the paradigms of *stas* and *tāns*, *ša-* in the forms which have *-a-* in the paradigm of *stas*, and *šie-* in the feminine. The initial *š-* must have spread from *ša-* to the other stems, as it has in Latvian.

The flexion of Li. *šīs* is identical with that of *jīs*: all case forms represent *šja-* with the exception of nom.sg. m. *šīs*, f. *ši*, acc.sg. m. *ši*, and loc.sg. m. *šimē* next to *šiamē* (cf. Daukša *jimē* next to *jamē*). Latvian has probably preserved an archaism in the vocalism of the masc. case forms: dat.sg. *šim*, acc.sg. *šuo*, loc.sg. *šimī*, *šinī*. If this is correct, the vocalism of the East Baltic masc. paradigm is identical with the one which I have proposed for its Prussian counterpart.

Putting the pieces together, we find that the Baltic and Slavic paradigms cannot be reduced to a single set of proto-forms. The Slavic paradigm appears to have originated from a conflation of the PIE.

particle *ki*, Balto-Slavic *śi*, with the paradigm of *e/i-*. The Baltic forms seem to have resulted from the addition of pronominal endings to the same particle, in Prussian apparently with generalization of *śī-* throughout the feminine. It is therefore probable that the form *śi* remained an uninflected particle up to the end of the Balto-Slavic period and was subsequently reshaped on the basis of the inflected pronouns in the separate languages.

3. We have now arrived at the reconstruction of a demonstrative pronoun *so*, *to-*, an anaphoric pronoun *e-*, *ei-*, *i-*, a suppletive nom.sg. form in *-n-*, and an uninflected particle *śi* meaning 'hic'. In addition, there is a Sl. stem *ovo-*, also meaning 'hic', correlating with *ono-* 'ille', and a stem *ino-*, meaning 'other' in Slavic and 'real' in Lithuanian and apparently related to PIE. *oino-* 'one'. The latter word will not be discussed here.

The stem *ovo-* is probably a Slavic derivative of the uninflected PIE. particle *au*, which is preserved in Gr. *αὖ* 'again', *αὐτός* 'self', Li. *aurè* 'there'. It preserved the original meaning 'that other' in Polish, Sorabian, and Slovene, whereas it replaced *s-* in Serbo-Croat and the contiguous South Slavic dialects (cf. Vaillant 1958: 380). If it can be suggested that *śi* and *au* were correlating particles in Balto-Slavic, designating 'hic' and 'istic' and used in combination with a neutral deictic stem, one may wonder if the suppletive nom.sg. form in *-n-* must be derived from a particle meaning 'illic'. This conjecture is indeed supported by the isolated Irish adverb *and* 'there', which is identical with Arm. *and* 'illic', cf. Li. *añdai* 'the other day', Gr. *ἔνῃ* 'the day after tomorrow',

ἐνθα 'there'. Assuming that Greek substituted the regular deictic element *e-* for *a-*, I reconstruct a Balto-Slavic particle *an* 'illic' which served as a basis for the creation of the 3rd person pronoun.

Starting from a single original demonstrative pronoun (*so, to-*), a single anaphoric pronoun (*e/i-*), and three deictic particles (*śi, au, an*), we can derive the Slavic and East Baltic paradigms in a straightforward way. The main problem which remains to be discussed is the rise of the Prussian forms. I think that the clue to a possible solution is provided by the Armenian system.

Unlike the other Indo-European languages, Armenian possesses three anaphoric pronouns, corresponding to the three demonstrative pronouns:

<i>ays</i> 'hic'	<i>sa</i> 'is'	<i>soyn</i> 'idem'
<i>ayd</i> 'iste'	<i>da</i> 'is'	<i>doyn</i> 'idem'
<i>ayn</i> 'ille'	<i>na</i> 'is'	<i>noyn</i> 'idem'

The radical element of the three series is identical with the one of Li. *šis, tās, anās*. As Holger Pedersen has pointed out (1982: 38f), the starting-point of the Armenian development must be sought in the forms *da*, which represents *to-*, and *ayn*, which is the reflex of *anio-*, Skt. *anyá-* 'other'. The latter word is a derivative of the proposed particle *an*, which has been preserved in the adverb *and* 'there', cf. also Skt. *ántara-* 'other', Li. *añtras* 'second'. Thus, we can derive the Arm. system from the one which has been reconstructed for Balto-Slavic.

The correctness of this analysis is corroborated by the forms *awasik* 'voici', *awadik*, *awanik* 'voilà'. Pedersen already identified *-sik* with Sl. *снѣ* as a derivative in *-ko-* of PIE. *ki* (1982: 41). It gave apparently rise to a correlating form *anik*, which in its

turn supplied the initial vowel for the creation of a form *asik*. Similarly, the initial vowel of *ast* 'here' was evidently taken from *and* 'there'. The initial part *aw-* may represent the PIE. particle *au*. Perikhanian's hypothesis that the forms under discussion are of Iranian origin does not explain the *-i-* (1966: 22). Of course, Iranian *hāu* is itself a conflation of *so* and *au*.

The close resemblance between the pre-Armenian and the reconstructed Balto-Slavic system and the presence of a threefold anaphoric pronoun in Armenian together offer an explanation for the origin of Pr. *tāns*. This pronoun evidently received its *t-* after the rise of a pronominal stem *ane-*, *anei-*, but before the contamination of *so* with *to-*. I therefore assume that there was a threefold anaphoric pronoun at an intermediate stage in the development of Prussian. The first step of the development was apparently the creation of a paradigm *an*, *ane-*, *anei-* 'he there'. This paradigm gave rise to the creation of correlating forms *te-*, *tei-* and *si-*, *siei-*. At the same time, or perhaps earlier, the demonstrative pronoun *so*, *sā*, *to-* gave rise to correlating paradigms *si*, *sio-*, fem. *sī*, *siā-*, and *ano-*, *anā-*. A form *tan* was probably created on the analogy of *an* in order to supply a nom.sg. form for the defective paradigm of *te-*, *tei-* because neither *so* nor *au* provided a suitable basis. When the threefold distinction in the anaphoric pronoun was lost, the stem *tan-* was generalized as a 3rd person pronoun. A trace of the earlier distribution is found in the substitutive use of *stas* for *tāns*, e.g. 75.13-22:

*Bhe kai prastan rettīwingiskan auskandinsnan en
tennesmu (adder stessie) auskiēndlai bhe semmai ēilai
wissan ka stesmu (adder tennēi) esse Adam engemmons ast
kai tans (adder tennā) dijgi is stesmu gīrbīn steison*

*nidruwīngin isklaitints ēnstan Swintan Arcan stēison
 Crixtiāniskun sausal bhe senpackai polaikūts Ainat
 garrewingī en Noseillien Wessals en podruwīsnan twaismu
 emnen schlūsīlai. Nostan kai tāns (adder tannā) sen
 wissamans Druwīngimans twaisei potaukinsnas Prabutskan
 gēiwan sengidaut musīlai "vnd das durch diese heilsame
 Sündflut an jm ersauffe vnd vntergehe alles was jm
 (oder jr) von Adam angeborn ist. Das er (oder sie) ausz
 der Zal der Vnglaubigen gesondert inn der heiligen
 Archa der Christenheit trucken vnd sicher behalten
 allezeit brünstig im Geist vnd frölich inn Hoffnung
 deinem Namen diene. Auff das er mit allen Glaubigen
 deiner Verheissung Ewiges leben erlangen moge".*

After the generalization of *tan-* in the anaphoric pronoun, the contamination of *so* and *to-* in the demonstrative pronoun yielded the regular paradigm of *stas*. This development may have started either in the nom.sg. form, where *so*, *sā* differed from *sī*, *sī* in the vowel only, or in the acc.sg. form, where *ton* and *tān* became *tan* as a result of the evolution of the vowel system, which created a disturbing homonymy with the pronominal stem *tan-*. In my view, this homonymy may have been the reason for the generalization of acc.sg. *-an*, pl. *-ans* in the forms *tennan*, *tennans*, as opposed to nom.pl. *tennei*, where the vocalism of the original anaphoric pronoun has been preserved. The resulting mixed paradigm provided in its turn the model for the substitution of *šī-* for *ša-* in the oblique cases of the demonstrative pronoun *schis*, which preserved the original vocalism in nom.pl. *schai*. These analogical changes obliterated the distinction between demonstrative and anaphoric pronouns, the only trace of which remains in the feminine gen. and dat.sg. forms of *stas*, where a consistent differentiation is made between the pronoun

(5x) and the article (15x).

4. The analysis proposed here offers an explanation for the formation of the Tocharian demonstrative pronoun. If we disregard B *samp* 'that there', which is obviously a compound of *se* with the adverb *omp* 'there', there are three paradigmatic sets: A *sām*, B *su* 'that', A *sās*, B *se* 'this', A *san*, B *sen* 'this here' (I write *-n* for the *anusvāra*). The inflected forms are the following:

- sg. nom. m. A *sām*, *sās*, *san*, B *su*, *se*, *sen*.
 f. A *sām*, *sās*, *sān*, B *sāu*, *sā*, *sān*.
 n. A *tām*, *tāṣ*, *tan*, B *tu*, *te*, *ten*.
 obl. m. A *cam*, *caṣ*, *can*, B *ceu*, *ce*, *cen*.
 f. A *tām*, *tāṣ*, *tān*, B *tāu*, *tā*, *tān*.
 pl. nom. m. A *cem*, *ceṣ*, *cen*, B *cey*, *cey*, *cey*.
 f. A *tom*, *toṣ*, *ton*, B *ton*, *toy*, *toyna*.
 obl. m. A *cesām*, *cesās*, *cesān*, B *cen*, *cen*, *ceyna*.
 f. A *tosām*, *tosās*, *tosān*, B *ton*, *toy*, *toyna*.

A comparative analysis leads to the following reconstruction of the Proto-Tocharian (PT.) forms:

- sg. nom. m. *se*, *sä-*, f. *sa*, n. *te*, *tä-*.
 obl. m. *ce*, f. *ta*.
 pl. nom. m. *cei*, f. *toü*.
 obl. m. *cens*, f. *tons*.

finals: *-u*, *-m*, *-š*, *-n*.

The extension of the final elements to the plural forms appears to be a recent development.

Lane has connected *-m* with the medial consonant of Sl. *romov* (1961: 475). This is doubtless correct. The other final elements must probably be identified with the three particles which have been reconstructed for Balto-Slavic. The expected reflex of PIE. *ki* is *-ś* rather than *-š*, but the alternation with *-s* in forms

with another *s* and the correspondence between A *näs* and B *ñäs* 'I, me', which can hardly be separated from the particle, suggest a neutralization of the opposition between the sibilants at the end of the word. I therefore assume that *-š* represents earlier *-ś*. The same particle is found as the 3rd sg. ending A *-s* in the present tense of the verb. Similarly, final *-n* is found in the corresponding B ending (cf. Pedersen 1941: 142f). An inflected form of the latter particle is attested in the 3rd sg. suffixed pronoun A *-n*, B *-ne*, suggesting an earlier paradigm which is strongly reminiscent of Balto-Slavic: nom. *an*, acc. *anom*.

The distribution of *-u*, *-m*, and *-š* does not require any special comments, but the semantic shift of *-n* is noteworthy. It can be compared with the South Slavic substitution of *ov-* for *s-* 'hic'. In the oldest texts, the Sl. stem is only used contrastively: *оѡѡ ... оѡѡ* or *оѡѡ .. ѡѡѡ*, cf. also Polish *ten ... ów*. Similarly, Toch. *-n* obtained the meaning 'hic' through its contrastive use in juxtaposition with the demonstrative pronoun. The original meaning was preserved in the suffixed anaphoric pronoun, and also in the indefinite pronoun A *san*, which is often found in combination with *ālak* 'other'. It is possible that B *omp* 'there' is related to Arm. *and* 'there'. The final *-p* may be identical with L1. *-p* 'at', Gr. *ἐπί* 'upon'.

The vocalism of the Toch. forms is remarkable. Since PT. *e* (A *a*, B *e*) is the phonetic reflex of PIE. *o*, B *se*, *te* developed regularly from *so*, *tod*. The same vowel is found before the final *-n*, but not before *-m*, *-š*, *-u*, where we find *sä-*, *tä-*, with *ä* representing PIE. *e*, *i*, but without the expected palatalization of the initial consonant before an original front vowel. It is therefore probable that we have to start from a

pronominal stem *e/i-* and that the addition of *s-*, *t-* was posterior to the palatalization.

The hypothesis that the pronominal stem *e/i-* was preserved in early Tocharian is corroborated by the palatalization of the initial consonant in PT. *ce* (A *ca-*, B *ce*), *cei* (A *ce-*, B *cey*), *cens* (A *ces-*, B *cen*). If we disregard the palatalization, these forms are the expected reflexes of *tom*, *toi*, *tons*. It follows that there must have been a stem *te-* with a suppletive nom.sg. form and that the initial consonant of this paradigm was adopted in the demonstrative pronoun. Here the Prussian forms come to mind. On the one hand, the absence of palatalization in nom.sg. *se*, *te* and its presence in *ce*, *cei*, *cens* is reminiscent of the reconstructed paradigm *tan*, *te-*, *tei-*. On the other, the discrepancy between the reduced vowel of nom.sg. *sä-*, *tä-* from *e-*, *i-* and the full vowel of *ce*, *cei*, *cens* from *to-* has its counterpart in Pr. *schis*, *schan*, *schai*, *schans*.

I now tentatively reconstruct a PT. anaphoric pronoun *äs*, *äu*, *än*, perhaps with an initial *y-*, obl. *cām* in noun phrases, *-ne* in verb phrases, and a demonstrative *se*, *sen*, obl. *ce*, *cen*, meaning 'this' when used contrastively with the anaphoric pronoun and designating simple deixis in juxtaposition with adverbial particles. The loss of the expected initial *y-* is also attested in B *ente*, A *äntāne* 'when' < PIE. *io-*, cf. B *intsu*, A *äntsān* 'which'. These forms point to a contamination of *io-*, *e/i-*, and *an*. The vocalism of the latter seems to be preserved in A *ān* 'whom', which is the phonetic reflex of *anom*.

Thus, we arrive at a reconstruction which strongly resembles the one proposed for Balto-Slavic. In both cases we have to start from a demonstrative pronoun

so, to-, an anaphoric pronoun e/i-, and three deictic particles *ki*, *au*, *an*. The attested paradigms originated from different kinds of contamination. One may therefore wonder if the feminine nom.pl. form PT. *toü* (A *to-*, B *toy*), which can be regarded as the phonetic reflex of *tāwes*, must be derived from a contamination of *tā-* and *au-*. I do not think that this is probable. The form is better connected with nom.pl. A *śnu* of *śän* 'wife', which points to the spread of *-wes* as a nom.pl. ending of feminine nouns at an early stage in the development of Tocharian.

University of Leiden

REFERENCES

- Kortlandt, F.H.H.
 1979 "On the history of the Slavic nasal vowels", *Indogermanische Forschungen* 84, 259-272.
 1982 "Innovations which betray archaisms", *Baltistica* 18/1, 4-9.
- Lane, G.S.
 1961 "On the formation of the Indo-European demonstrative", *Language* 37, 469-475.
- Pedersen, H.
 1941 *Tocharisch vom Gesichtspunkt der indoeuropäischen Sprachvergleichung*, København.
 1982 *Kleine Schriften zum Armenischen*, Hildesheim.
- Perikhanian, A.
 1966 "Une inscription araméenne du roi Artasēs trouvée à Zanguézour", *Revue des études arméniennes*, NS. 3, 17-29.
- Vaillant, A.
 1958 *Grammaire comparée des langues slaves II: Morphologie*, Lyon.
- Van Wijk, N.
 1918 *Altpreussische Studien*, Haag.